Metrika (2013) 76:577–593
DOI 10.1007/s00184-012-0405-9

On optimal choice of order statistics in large samples for the construction of confidence regions for the location and scale

Alexander Zaigraev · Magdalena Alama-Bućko

Received: 20 March 2011 / Published online: 5 August 2012

Abstract Given a large sample from a location-scale population we estimate the unknown parameters by means of confidence regions constructed on the basis of two order statistics. The problem of the best choice of those statistics to obtain good estimates, as $n \to \infty$, is considered.

Keywords Order statistics · Optimal confidence regions · Pivot · Location-scale population

1 Introduction

A problem of optimal choice of order statistics in large samples for the best estimation of the location and scale is not new. For example, Subsection 10.4 of David and Nagaraja (2003) is devoted to such a problem in case of the point estimation (see also the references cited therein). However, the same problem for the confidence region estimation has not attracted the attention so far, as far as we know. This paper is an attempt to fill the gap.

Let $x = (x_1, x_2, \ldots, x_n)$ be a sample from a distribution P_θ, $\theta = (\theta_1, \theta_2)$, that is $\{x_i\}$ are independent real-valued random variables having the distribution P_θ. We deal with the case where $\theta_1 \in R$ is a location parameter and $\theta_2 > 0$ is a scale parameter. As the estimators of $\theta = (\theta_1, \theta_2)$, let us consider two-dimensional confidence regions.

A. Zaigraev (✉)
Faculty of Mathematics and Computer Science, Nicolaus Copernicus University,
Chopin str. 12/18, 87-100, Toruń, Poland
e-mail: alzaig@mat.umk.pl

M. Alama-Bućko
Institute of Mathematics and Physics, University of Technology and Life Sciences,
ul. Kaliskiego 7, 85-796, Bydgoszcz, Poland
e-mail: mbucko@utp.edu.pl

Let $\alpha \in (0, 1)$ be a given confidence level. A strong confidence region of level α is a mapping $B : R^n \mapsto \mathcal{B}^2$ such that

$$P_\theta(\theta \in B(x)) = \alpha \quad \forall\theta,$$

where $\mathcal{B}^2$ is the σ-algebra of Borel subsets of R^2. The quality of a confidence region can be characterized by the risk function defined as

$$R(\theta, B) = E_\theta \lambda_2(B(x)),$$

where λ_2 is the Lebesgue measure on $\mathcal{B}^2$. Among strong confidence regions we distinguish those having the minimal risk and call them optimal.

The method for construction of an optimal confidence region is well-known (see, for example, Alama-Bućko et al. 2006 or Czarnowska and Nagaev 2001) and is based on using a pivot. Let $t_1(x)$ and $t_2(x)$ be a couple of statistics satisfying the following conditions: for any $a \in R,\ b > 0$,

$$t_1(bx + a\mathbf{1}_n) = bt_1(x) + a, \qquad t_2(bx + a\mathbf{1}_n) = bt_2(x), \tag{1}$$

where $\mathbf{1}_n = (1, 1, \ldots, 1) \in R^n$. Let from now on $y = (y_1, y_2, \ldots, y_n)$ be a sample from the standard distribution $P_{(0,1)}$. Taking a set $A \in \mathcal{B}^2$ such that

$$P_{(0,1)}\left(\left(-\frac{t_1(y)}{t_2(y)}, \frac{1}{t_2(y)} - 1\right) \in A\right) = \alpha, \tag{2}$$

one can obtain due to (1)

$$P_{(\theta_1,\theta_2)}\left(\left(\frac{\theta_1 - t_1(x)}{t_2(x)}, \frac{\theta_2 - t_2(x)}{t_2(x)}\right) \in A\right) = \alpha. \tag{3}$$

That is,

$$\left(\frac{\theta_1 - t_1(x)}{t_2(x)}, \frac{\theta_2 - t_2(x)}{t_2(x)}\right) \tag{4}$$

is a pivot.

Thus, the set

$$B_A(x) = (t_1(x), t_2(x)) + t_2(x)A \tag{5}$$

is a strong confidence region for (θ_1, θ_2). In this case,

$$R(\theta, B_A) = \lambda_2(A)E_\theta t_2^2(x) = \theta_2^2 \lambda_2(A)E_{(0,1)}t_2^2(y), \tag{6}$$

that is the risk function is proportional to the area of the set A, and the problem is to choose the set A with the smallest area.

Assume that the density function g of the random vector $(-t_1(y)/t_2(y), 1/t_2(y)-1)$ exists, continuous and such that

$$\lambda_2(\{u \in R^2 : \ g(u) = z\}) = 0 \quad \forall z \geq 0.$$

The confidence region is optimal among all the confidence regions of the form (5), if

$$A = \{u \in R^2 : \ g(u) \geq z_\alpha\},$$

where z_α is defined by the equation

$$\int_A g(u)\mathrm{d}u = \alpha.$$

This is a corollary of Proposition 2.1 of Einmahl and Mason (1992).

Of course, the optimal confidence region depends on the choice of t_1 and t_2.

For the natural interpretation of confidence region (5) it is reasonable to take as $t_1(x)$ and $t_2(x)$ the estimators of the location and scale parameters, respectively. Then $(t_1(x), t_2(x))$ is the center of the region, while the set A defines the shape of the region and $t_2(x)$ is responsible for its rescaling.

In this paper we consider the case, where t_1 and t_2 are linear functions of two order statistics. Some other cases were considered in Alama-Bućko et al. (2006) and Czarnowska and Nagaev (2001).

Let $x_{k:n}$ and $x_{m:n}$ be the k-th and the m-th order statistic of the sample x, respectively, $k < m$. The main goal of the paper is to make the best possible choice of $k = k_n$ and $m = m_n$ to minimize risk function (6), as $n \to \infty$, under the assumption that $k/n \to p, \ m/n \to q, \ p < q$.

Asymptotics of the optimal confidence region in case $0 < p < q < 1$ is obtained in Sect. 2. Our main results are established in Sect. 3, while Sect. 4 contains examples. In "Appendix" we prove three useful auxiliary lemmas.

2 Asymptotics of the optimal confidence region

Let $F = F_{(0,1)}$ be the continuous distribution function corresponding to $P_{(0,1)}$ and

$$F^{-1}(p) = \inf\{t \in R : \ F(t) \geq p\}, \quad 0 < p < 1$$

be the so-called quantile function. We assume that the distribution F is absolutely continuous and denote by f its density function. Let φ_V be the density corresponding to the normal distribution with zero mean vector and covariance matrix V.

We start with the classical result on limit distribution for central order statistics (see, for example, Theorem 10.3 of David and Nagaraja 2003 or Theorem 4.1.3 of Reiss 1989).

Proposition 1 *Let* $0 < p < q < 1$ *be fixed and* $k/n - p = o(n^{-1/2}), \ m/n - q = o(n^{-1/2})$, *as* $n \to \infty$. *Assume also that* $f(F^{-1}(p)) > 0, \ f(F^{-1}(q)) > 0$. *Then the limit distribution of the vector* $n^{1/2}(y_{k:n} - F^{-1}(p), y_{m:n} - F^{-1}(q))$, *as* $n \to \infty$, *is normal with zero mean vector and covariance matrix*

$$V = \begin{pmatrix} \dfrac{p(1-p)}{f^2(F^{-1}(p))} & \dfrac{p(1-q)}{f(F^{-1}(p))f(F^{-1}(q))} \\ \dfrac{p(1-q)}{f(F^{-1}(p))f(F^{-1}(q))} & \dfrac{q(1-q)}{f^2(F^{-1}(q))} \end{pmatrix}.$$

Assume that $F^{-1}(q) \neq F^{-1}(p)$ and take

$$t_1(x) = \frac{x_{k:n}F^{-1}(q) - x_{m:n}F^{-1}(p)}{F^{-1}(q) - F^{-1}(p)}, \qquad t_2(x) = \frac{x_{m:n} - x_{k:n}}{F^{-1}(q) - F^{-1}(p)}. \tag{7}$$

Note that $t_1(x)$ and $t_2(x)$ from (7) satisfy (1) and are asymptotically unbiased estimators of the location and scale, respectively.

Making use of statement (i) of Lemma 1 from "Appendix" with $a_n = c_n = n^{1/2}$, $b_n = F^{-1}(p)$, $d_n = F^{-1}(q)$, $\xi_n = y_{k:n}$, $\eta_n = y_{m:n}$, $f(u_1, u_2) = \varphi_V(u_1, u_2)$, we immediately obtain the following result.

Corollary 1 *Under the conditions of Proposition 1, the limit distribution of the vector* $n^{1/2}(-t_1(y)/t_2(y), 1/t_2(y) - 1)$, *where* t_1 *and* t_2 *are defined by* (7)*, as* $n \to \infty$, *is normal with zero mean vector and the covariance matrix* $W = (H^{-1})^T V H^{-1}$, *where*

$$H = \begin{pmatrix} 1 & 1 \\ F^{-1}(p) & F^{-1}(q) \end{pmatrix}.$$

Applying the method of construction of optimal confidence regions described in Sect. 1 (see formulae (2)–(5)), one can obtain the following optimal confidence region based on the vector $n^{1/2}(-t_1(y)/t_2(y), 1/t_2(y) - 1)$:

$$B_{A_n}(x) = (t_1(x), t_2(x)) + (t_2(x)/\sqrt{n})A_n, \tag{8}$$

where the set A_n is defined by

$$A_n = \{u \in R^2 : \ g_n(u) \geq z_\alpha\}, \int\limits_{A_n} g_n(u)\mathrm{d}u = \alpha,$$

and g_n is the density corresponding to $n^{1/2}(-t_1(y)/t_2(y), 1/t_2(y) - 1)$. The corresponding risk function has the form

$$R(\theta, B_{A_n}) = E_\theta t_2^2(x)\lambda_2(A_n)/n. \tag{9}$$

Let us investigate the behaviour of $R(\theta, B_{A_n})$ as $n \to \infty$. From Proposition 1 it follows that

$$\lim_{n\to\infty} E_\theta t_2^2(x) = \theta_2^2 \lim_{n\to\infty} E_{(0,1)} t_2^2(y) = \theta_2^2. \tag{10}$$

Moreover, basing on Proposition 1, as it was shown in Theorem 2 of Alama-Bućko and Zaigraev (2006), one can obtain the asymptotic expansion of the set A_n as $n \to \infty$. Namely, the set A_n, as $n \to \infty$, approximates the ellipse A_0 of the form

$$A_0 = \{u \in R^2 : \varphi_W(u) \geq z'_\alpha\},$$

where z'_α is defined by the equation

$$\int\limits_{A_0} \varphi_W(u)\mathrm{d}u = \alpha, \quad \text{that is} \quad z'_\alpha = \frac{1-\alpha}{2\pi(\det W)^{1/2}}.$$

In other words,

$$A_0 = \left\{u \in R^2 : \frac{e^{-uW^{-1}u^T/2}}{2\pi(\det W)^{1/2}} \geq z'_\alpha\right\},$$

or

$$A_0 = \{u \in R^2 : uW^{-1}u^T \leq -2\ln(1-\alpha)\}.$$

Therefore,

$$\lim_{n\to\infty} \lambda_2(A_n) = \lambda_2(A_0) = -2\pi \ln(1-\alpha)(\det W)^{1/2}$$

and

$$(\det W)^{1/2} = \frac{(\det V)^{1/2}}{\det H} = \frac{[p(q-p)(1-q)]^{1/2}}{[F^{-1}(q) - F^{-1}(p)]f(F^{-1}(p))f(F^{-1}(q))}. \tag{11}$$

Summing up, $R(\theta, B_{A_n})$ is of order $1/n$ as $n \to \infty$, if $0 < p < q < 1$, $F^{-1}(q) \neq F^{-1}(p)$, $f(F^{-1}(q)) > 0$, $f(F^{-1}(p)) > 0$.

The problem of interest is to search for p^* and q^* to minimize (11). In other words, this is the problem of choice the order statistics $x_{k:n}, x_{m:n}$ to obtain the optimal confidence region for θ with the smallest risk function.

3 Optimal choice of order statistics

After changing the notation $u = F^{-1}(p)$, $v = F^{-1}(q)$, the right-hand side of (11) is rewritten as

$$\frac{[F(u)(F(v) - F(u))(1 - F(v))]^{1/2}}{(v-u)f(u)f(v)} := G(u, v).$$

Let

$$(u^*, v^*) \in \arg \inf_{u_F^- < u < v < u_F^+} G(u, v),$$

where $-\infty \le u_F^- < u_F^+ \le \infty$ are the lower and the upper end of the support of the distribution F, respectively.

Note that for any fixed $u \in (u_F^-, u_F^+)$,

$$\lim_{v \downarrow u} (v-u)^{1/2} G(u,v) = \frac{[F(u)(1-F(u))]^{1/2}}{(f(u))^{3/2}}.$$

Therefore, $G(u,v) \uparrow \infty$ as $v \downarrow u$ for any fixed $u \in (u_F^-, u_F^+)$. Similarly, it can be shown that $G(u,v) \uparrow \infty$ as $u \uparrow v$ for any fixed $v \in (u_F^-, u_F^+)$.

In what follows, we assume that the function f is differentiable at any point $u \in (u_F^-, u_F^+)$. Then

$$(\ln(G(u,v))'_v = -\frac{f'(v)}{f(v)} - \frac{1}{v-u} + \frac{f(v)}{2}\left(\frac{1}{F(v)-F(u)} - \frac{1}{1-F(v)}\right) := H(u,v). \tag{12}$$

By simple calculations,

$$\lim_{v \downarrow u} (v-u) H(u,v) = -\frac{1}{2}$$

and, therefore, $H(u,v) < 0$ in a neighborhood of any fixed $u \in (u_F^-, u_F^+)$, that is the function $G(u,v)$ is decreasing in v in a neighborhood of u. Similarly, the function $G(u,v)$ is increasing in u in a neighborhood of any fixed $v \in (u_F^-, u_F^+)$ since

$$(\ln(G(u,v))'_u = -\frac{f'(u)}{f(u)} + \frac{1}{v-u} + \frac{f(u)}{2}\left(\frac{1}{F(u)} - \frac{1}{F(v)-F(u)}\right) := H^*(u,v) \tag{13}$$

and

$$\lim_{u \uparrow v} (v-u) H^*(u,v) = \frac{1}{2}.$$

In the sequel we need some well-known facts from the extreme value theory (see, for example, Subsection 10.5 of David and Nagaraja 2003).

If there exist $c_n > 0$ and $d_n \in R$ such that the limit distribution of the sequence $c_n(y_{n:n} - d_n)$ exists, as $n \to \infty$, then the limit distribution function is one of just three types ($\beta > 0$):

– (Fréchet) $H_1(u;\beta) = \begin{cases} 0, & u \le 0 \\ \exp(-u^{-\beta}), & u > 0, \end{cases}$

– (Weibull) $H_2(u;\beta) = \begin{cases} \exp(-(-u)^{\beta}), & u \le 0 \\ 1, & u > 0, \end{cases}$

– (Gumbel) $H_3(u) = \exp(-\exp(-u)),\ u \in R.$

In this case it is said that F belongs to the domain of attraction of the distribution $H_i, i = 1,2,3$ (written $F \in D(H_i)$).

Let h be the hazard rate function, that is

$$h(u) = \frac{f(u)}{1 - F(u)}, \quad u \in (u_F^-, u_F^+).$$

It turns out that the possible limit laws for the properly centered and normed maximal order statistics $x_{n:n}$ are determined by the behaviour of the function h in a neighborhood of the right endpoint of F. The following result (see, for example, Theorems 8.3.3 and 8.3.4 of Arnold et al. 1992) contains the well-known sufficient von Mises conditions of attraction to $D(H_i)$, $i = 1, 2, 3$, and description of sequences $\{c_n, d_n\}$. In what follows, $L(v)$ denotes a slowly varying function as $v \to \infty$.

Proposition 2 *The following statements hold:*

– $F \in D(H_1)$, *if* $u_F^+ = \infty$ *and for some* $\beta > 0$,

$$\lim_{u \to \infty} uh(u) = \beta; \tag{14}$$

here $d_n = 0$, $c_n = (F^{-1}(1 - 1/n))^{-1} = n^{-1/\beta} L(n)$;
– $F \in D(H_2)$, *if* $u_F^+ < \infty$ *and for some* $\beta > 0$,

$$\lim_{u \to u_F^+} (u_F^+ - u)h(u) = \beta; \tag{15}$$

here $d_n = u_F^+$, $c_n = (u_F^+ - F^{-1}(1 - 1/n))^{-1} = n^{1/\beta} L(n)$;
– $F \in D(H_3)$, *if* $f(u)$ *is differentiable for all* $u > u_0$ *and*

$$\lim_{u \to u_F^+} (1/h(u))' = 0; \tag{16}$$

here $d_n = F^{-1}(1 - 1/n)$, $c_n = h(d_n) = nf(d_n)$.

Remark 1 Comparing the norming sequences $\{c_n\}$ from all the above cases to $n^{1/2}$, one can conclude that: $c_n \gg n^{1/2}$ if $F \in D(H_2)$ with $\beta < 2$, while $c_n \ll n^{1/2}$ if $F \in D(H_1)$, or $F \in D(H_3)$, or $F \in D(H_2)$ with $\beta > 2$. In what follows, we exclude the case $F \in D(H_2)$ with $\beta = 2$ from the consideration since uncertainty remains here.

Now we are able to establish the crucial result for optimal choice of order statistics.

Theorem 1 *The following statements hold for any fixed* $u \in (u_F^-, u_F^+)$:

(i) *if condition* (14) *holds, then* $G(u, v) \uparrow \infty$ *as* $v \uparrow u_F^+ = \infty$;
(ii) *if condition* (16) *holds, then* $G(u, v)$ *is a non-decreasing function for all* $v > v_0$;
(iii) *if condition* (15) *holds, then*

$$\beta > 2 \implies G(u, v) \uparrow \infty, \ v \uparrow u_F^+,$$
$$\beta < 2 \implies G(u, v) \downarrow 0, \ v \uparrow u_F^+.$$

Proof Statement (i) is a direct consequence of (14).

To prove the statement (ii), it is enough to show that

$$\lim_{v\uparrow u_F^+} \frac{H(u,v)}{h(v)} > 0. \tag{17}$$

From (12) it follows that

$$\frac{H(u,v)}{h(v)} = \left(\frac{1}{h(v)}\right)' - \frac{1}{(v-u)h(v)} + \frac{1}{2}\cdot\frac{1-F(u)}{1-F(u)-[1-F(v)]}.$$

In view of condition (16), it is enough to prove that

$$\lim_{v\uparrow u_F^+} \frac{1}{(v-u)h(v)} = 0 \qquad \left(\text{then } \lim_{v\uparrow u_F^+} \frac{H(u,v)}{h(v)} = \frac{1}{2}\right). \tag{18}$$

For this purpose one can use the arguments from the proof of Remark 2 of Subsection 3.3.3 of Embrechts et al. (1997). If $u_F^+ = \infty$, then since $(1/h(v))' \downarrow 0$, as $v \uparrow \infty$, the Cesàro mean of this function also converges, that is

$$\lim_{v\uparrow\infty} \frac{1/h(v)}{v} = \lim_{v\uparrow\infty} \frac{1}{v}\int_z^v (1/h(u))'du = 0.$$

If $u_F^+ < \infty$, then

$$\lim_{v\uparrow u_F^+} \frac{1/h(v)}{u_F^+ - v} = -\lim_{v\uparrow u_F^+} \int_v^{u_F^+} \frac{(1/h(u))'du}{u_F^+ - v} = \lim_{t\downarrow 0} \frac{1}{t}\int_0^t (1/h(u_F^+ - s))'ds.$$

Since $(1/h(u_F^+ - s))' \downarrow 0$, as $s \downarrow 0$, the last limit tends to 0 and (18) holds.

Statement (iii) follows from Theorem 3.3.12 of Embrechts et al. (1997) and properties of slowly varying functions.

Theorem 1 immediately implies the following result.

Corollary 2 *If condition* (15) *holds with* $\beta < 2$, *then* $v^* = u_F^+ < \infty$ $(q^* = 1)$ *and* $\inf G(u,v) = 0$. *In other cases (condition* (14)*, or condition* (16)*, or condition* (15) *with* $\beta > 2$ *holds),* $v^* < u_F^+$ $(q^* < 1)$ *and* $\inf G(u,v) > 0$.

As it is known, similar results hold also for the minimal order statistic $y_{1:n}$. More precisely, if there exist $a_n > 0$ and $b_n \in R$ such that the limit distribution of the sequence $a_n(y_{1:n} - b_n)$ exists, as $n \to \infty$, then the limit distribution function is one of just three types: $H_i^*(u;\gamma) = 1 - H_i(-u;\gamma)$, $i = 1, 2, 3$. So, with the small evident modifications one can establish for $y_{1:n}$ the similar results as for $y_{n:n}$. We have gathered them in the following theorem.

Denote

$$h^*(u) = \frac{f(u)}{F(u)}, \quad u \in (u_F^-, u_F^+).$$

Theorem 2 *The following statements hold for any fixed* $v \in (u_F^-, u_F^+)$.

1. *If* $u_F^- = -\infty$ *and for some* $\gamma > 0$,

$$\lim_{u \to -\infty} u h^*(u) = -\gamma, \tag{19}$$

then $F \in D(H_1^*)$ *and* $G(u, v) \uparrow \infty$ *as* $u \downarrow u_F^- = -\infty$.

2. *If* $f(u)$ *is differentiable for all* $u < u_1^*$ *and*

$$\lim_{u \to u_F^-} (1/h^*(u))' = 0, \tag{20}$$

then $F \in D(H_3^*)$ *and* $G(u, v)$ *is a non-increasing function for* $u < u_0^*$.

3. *If* $u_F^- > -\infty$ *and for some* $\gamma > 0$,

$$\lim_{u \to u_F^-} (u - u_F^-) h^*(u) = \gamma, \tag{21}$$

then $F \in D(H_2^*)$ *and*

$$\gamma > 2 \implies G(u, v) \uparrow \infty, \; u \downarrow u_F^-,$$
$$\gamma < 2 \implies G(u, v) \downarrow 0, \; u \downarrow u_F^-.$$

If condition (21) *with* $\gamma < 2$ *holds, then* $u^* = u_F^- > -\infty$ $(p^* = 0)$ *and* $\inf G(u, v) = 0$. *In other cases (condition* (19)*, or condition* (20)*, or condition* (21) *with* $\gamma > 2$ *holds),* $u^* > u_F^-$ $(p^* > 0)$ *and* $\inf G(u, v) > 0$.

Summing up, assuming that the underlying distribution in a neighborhood of u_F^+ satisfies one of von Mises conditions (14)–(16) and in a neighborhood of u_F^- satisfies one of von Mises conditions (19)–(21), we can formulate the results on optimal choice of order statistics distinguishing between four cases.

Case I If in a neighborhood of u_F^+ (14), or (16), or (15) with $\beta > 2$ holds and in a neighborhood of u_F^- (19), or (20), or (21) with $\gamma > 2$ holds, then $0 < p^* < q^* < 1$. In this case we take (see (7))

$$t_1(x) = \frac{x_{k^*:n} F^{-1}(q^*) - x_{m^*:n} F^{-1}(p^*)}{F^{-1}(q^*) - F^{-1}(p^*)}, \quad t_2(x) = \frac{x_{m^*:n} - x_{k^*:n}}{F^{-1}(q^*) - F^{-1}(p^*)}, \tag{22}$$

where, for example, $k^* = [np^*] + 1, \; m^* = [nq^*] + 1$. The optimal confidence region is based on the vector $T_n(y) = n^{1/2}(-t_1(y)/t_2(y), 1/t_2(y) - 1)$ and is given by (8) with risk function (9), where $\lim_{n \to \infty} E_\theta t_2^2(x) = \theta_2^2$ (see (10)). The limit law for $T_n(y)$, established in Corollary 1, allows us to state that

$$\lambda_2(A_n) \to (-2\pi \ln(1-\alpha)) \frac{[p^*(q^* - p^*)(1-q^*)]^{1/2}}{[F^{-1}(q^*) - F^{-1}(p^*)] f(F^{-1}(p^*)) f(F^{-1}(q^*))}, \quad n \to \infty.$$

The risk function is of order $1/n$, as $n \to \infty$.

The important note: the order of the risk function for the optimal confidence region equals to the reciprocal of the product of norming sequences of the components of $T_n(y)$, that is $1/n = 1/(n^{1/2} \cdot n^{1/2})$.

It remains to consider the cases when in a neighborhood of u_F^+ ($u_F^+ < \infty$) (15) with $\beta < 2$ holds and/or in a neighborhood of u_F^- ($u_F^- > -\infty$) (21) with $\gamma < 2$ holds. Here, the order of the corresponding risk function is evidently $o(1/n)$ and $q^* = 1$ and/or $p^* = 0$. In this case we need to change the vector $T_n(y)$ and norming sequences according to statements (ii) or (iii) of Lemma 1 from "Appendix". Again the reciprocal of the product of norming sequences of the components of $T_n(y)$ determines the order of the risk function.

Let $q^* = 1$ (the case $p^* = 0$ can be considered similarly). In general, one can distinguish between three types of sequences $\{m_n\}$ satisfying $m_n/n \to q^* = 1, n \to \infty$:

(a) $m_n = n$; in this case $y_{m_n:n} = y_{n:n}$, i. e. we deal with the extreme order statistics;
(b) $m_n = n - j + 1,\ j > 1$ is fixed; in this case $y_{m_n:n} = y_{n-j+1:n}$, i. e. we deal with *other extreme order statistics*;
(c) $m_n = n - j + 1,\ j = j_n \to \infty,\ j_n/n \to 0,\ n \to \infty$; in this case $y_{m_n:n} = y_{n-j_n+1:n}$, i. e. we deal with *intermediate order statistics*.

The question arises: what type of the sequence one should choose to obtain the better confidence region?

First of all, note that according to the end of Subsection 10.8 of David and Nagaraja (2003), lower extremes are asymptotically independent of upper extremes and both are asymptotically independent of central order statistics as well as of intermediate order statistics.

In situation (a) the possible limit laws and corresponding norming sequences are given in Proposition 2. In situation (b), as it follows from Theorem 8.4.1 of Arnold et al. (1992), $F \in D(H_i),\ i = 1, 2, 3$, iff the limit distribution function of an extreme order statistic $y_{n-j+1:n}$, as $n \to \infty$, where j is fixed, is of the form $\sum_{r=0}^{j-1} H_i(u)[-\ln(H_i(u))]^r/r!$, $i = 1, 2, 3$; the sequences $\{c_n, d_n\}$ are the same as in Proposition 2. Therefore, comparing the choice of $y_{n:n}$ with that of $y_{n-j+1:n}$, where $j > 1$ is fixed, we conclude that the norming sequences are the same, but the first choice is better since it gives the shorter interval for the appropriate coordinate (see Lemma 2 from "Appendix").

At last, in situation (c), as it follows from Theorem 8.5.3 of Arnold et al. (1992), if von Mises conditions (14)–(16) hold, then the limit law for $y_{n-j+1:n},\ n \to \infty,\ j \to \infty,\ j/n \to 0$, is standard normal and $d_n = F^{-1}(1 - j/n),\ c_n = nf(d_n)/j^{1/2}$. Note that in the case of interest (when in a neighborhood of u_F^+ (15) with $\beta < 2$ holds) this norming sequence $\{c_n\}$ is less than that for $y_{n:n}$ given in Proposition 2 (see Lemma 3 from "Appendix"); in all other cases it is less than $n^{1/2}$.

Case II If in a neighborhood of u_F^+ (one can take $u_F^+ = 0$ without loss in generality) (15) with $\beta < 2$ holds, while in a neighborhood of u_F^- (19), or (20), or (21) with $\gamma > 2$

holds, then $0 < p^* < q^* = 1$. In this case, drawing on (22), we take

$$t_1(x) = x_{n:n}, \qquad t_2(x) = -\frac{x_{n:n} - x_{k^*:n}}{F^{-1}(p^*)}.$$

The pivotal quantity is

$$\left(c_n \frac{\theta_1 - t_1(x)}{t_2(x)}, n^{1/2} \frac{\theta_2 - t_2(x)}{t_2(x)}\right),$$

where $c_n = n^{1/\beta} L(n) \gg n^{1/2},\ n \to \infty$. The optimal confidence region is based on the vector $(-c_n t_1(y)/t_2(y), n^{1/2}(1/t_2(y) - 1))$ and has the risk of order $1/(c_n n^{1/2}) \ll 1/n$, as $n \to \infty$. It has the form

$$B_{A_n^*}(x) = (t_1(x), t_2(x)) + t_2(x) A_n^*,$$

where $A_n^* = \{(z_1, z_2) :\ (c_n z_1, n^{1/2} z_2) \in A_n\}$.

Case III If in a neighborhood of u_F^+ (14), or (16), or (15) with $\beta > 2$ holds, while in a neighborhood of u_F^- (one can take $u_F^- = 0$ without loss in generality) (21) with $\gamma < 2$ holds, then $0 = p^* < q^* < 1$. In this case, drawing on (22), we take

$$t_1(x) = x_{1:n}, \qquad t_2(x) = \frac{x_{m^*:n} - x_{1:n}}{F^{-1}(q^*)}.$$

The pivotal quantity is

$$\left(a_n \frac{\theta_1 - t_1(x)}{t_2(x)}, n^{1/2} \frac{\theta_2 - t_2(x)}{t_2(x)}\right),$$

where $a_n = n^{1/\gamma} L(n) \gg n^{1/2},\ n \to \infty$. The optimal confidence region is based on the vector $(-a_n t_1(y)/t_2(y), n^{1/2}(1/t_2(y) - 1))$ and has the risk of order $1/(a_n n^{1/2}) \ll 1/n$, as $n \to \infty$. It has the form

$$B_{A_n^*}(x) = (t_1(x), t_2(x)) + t_2(x) A_n^*,$$

where $A_n^* = \{(z_1, z_2) :\ (a_n z_1, n^{1/2} z_2) \in A_n\}$.

Case IV At last, if in a neighborhood of u_F^+ (15) with $\beta < 2$ holds and in a neighborhood of u_F^- (21) with $\gamma < 2$ holds, then $p^* = 0,\ q^* = 1$. In this case the construction repeats one of the previous cases depending on the relation between β and γ (see Examples).

At last, it is worth to note that if the distribution F is symmetric, that is its density f satisfies the condition $f(-u) = f(u)$, and, moreover, if f is a differentiable infinitely many times function such that $f'(-u) = -f'(u)$, then $p^* = 1 - q^*$ (see Theorem 10.4 of David and Nagaraja 2003 and also Ogawa 1998 for the proof and discussion).

4 Examples

Here we consider three examples of distributions. In all the cases we calculate the values of the risk function, according to (6): firstly, for $(p, q) = (0.25, 0.75)$, and secondly, for (p^*, q^*). Even for the realistic sample sizes, the risk is smaller in the second case.

Example 1 Uniform distribution $U(\theta_1 - \theta_2/2, \theta_1 + \theta_2/2)$.

It is the case IV since in a neighborhood of $u_F^+ = 1/2$ (15) with $\beta = 1$ holds, while in a neighborhood of $u_F^- = -1/2$ (21) with $\gamma = 1$ holds. Therefore, $(p^*, q^*) = (0, 1)$. Here $\beta = \gamma$, and the optimal confidence region for (θ_1, θ_2) is based on

$$t_1(x) = (x_{1:n} + x_{n:n})/2, \qquad t_2(x) = x_{n:n} - x_{1:n}$$

while the pivot looks as $n(\frac{\theta_1 - t_1(x)}{t_2(x)}, \frac{\theta_2}{t_2(x)} - 1)$.

The optimal confidence region has the risk of order $1/n^2$.

Calculations of the risk function for $(p, q) = (0.25, 0.75)$ (the first table) and for $(p^*, q^*) = (0, 1)$ (the second table):

n	k	m	$\lambda_2(A)$	$\mathrm{E}_{(0,1)} t_2^2(y)$	$R(\theta, B_A)/\theta_2^2$
30	8	23	1.076282	0.241935	0.260390
40	10	31	0.640971	0.243902	0.156334
50	13	38	0.599687	0.245098	0.146982
60	15	46	0.429006	0.245901	0.105493
70	18	53	0.414028	0.246478	0.102049
80	20	61	0.326711	0.246913	0.080669
90	23	68	0.312016	0.247252	0.077147
100	25	76	0.262212	0.247524	0.064904
500	125	376	0.052771	0.249500	0.013166

n	k	m	$\lambda_2(A)$	$\mathrm{E}_{(0,1)} t_2^2(y)$	$R(\theta, B_A)/\theta_2^2$
30	1	30	0.015230	0.877016	0.013357
40	1	40	0.008145	0.905923	0.007379
50	1	50	0.005059	0.923831	0.004674
60	1	60	0.003444	0.936012	0.003224
70	1	70	0.002495	0.944835	0.002357
80	1	80	0.001890	0.951520	0.001798
90	1	90	0.001481	0.956760	0.001417
100	1	100	0.001192	0.960978	0.001145
500	1	500	0.000050	0.992039	0.000049

Example 2 Exponential distribution $E(\theta_1, \theta_2)$.

It is the case III since in a neighborhood of $u_F^- = 0$ (21) with $\gamma = 1$ holds, while in a neighborhood of $u_F^+ = \infty$ (16) holds.

By simple calculations we obtain

$$G(u, v) = \frac{[(e^{2u} - e^u)(e^{v-u} - 1)]^{1/2}}{v - u}, \quad 0 < u < v < \infty,$$

and $\arg \inf_{0<u<v<\infty} G(u, v) = (0, v_*)$, where v_* is the solution of the equation $(1 - v/2)e^v = 1$ ($v_* = 1.5936$). Therefore, $(p^*, q^*) = (0, v_*/2) = (0, 0.7968)$.

The optimal confidence region has the risk of order $1/n^{3/2}$.

Calculations of the risk function for $(p, q) = (0.25, 0.75)$ (the first table) and for $(p^*, q^*) = (0, 0.7968)$ (the second table):

n	k	m	$\lambda_2(A)$	$\mathrm{E}_{(0,1)} t_2^2(y)$	$R(\theta, B_A)/\theta_2^2$
30	8	23	0.636060	1.294206	0.823192
40	10	31	0.361046	1.431981	0.517011
50	13	38	0.332177	1.259720	0.418450
60	15	46	0.237054	1.354291	0.321040
70	18	53	0.224124	1.244763	0.278981
80	20	61	0.177264	1.316461	0.233361
90	23	68	0.171444	1.236410	0.211975
100	25	76	0.141810	1.294084	0.183514
500	125	376	0.027746	1.224075	0.033963

n	j	k	$\lambda_2(A)$	$\mathrm{E}_{(0,1)} t_2^2(y)$	$R(\theta, B_A)/\theta_2^2$
30	1	24	0.059575	2.783466	0.165825
40	1	32	0.036114	2.973219	0.107374
50	1	40	0.025097	3.149164	0.079034
60	1	48	0.018263	3.320204	0.060636
70	1	56	0.014387	3.490169	0.050213
80	1	64	0.011406	3.660958	0.041756
90	1	72	0.009464	3.833603	0.036281
100	1	80	0.008152	4.008691	0.032678
500	1	399	0.000517	12.822575	0.006629

Example 3 Normal distribution $N(\theta_1, \theta_2)$.

It is the case I since in a neighborhood of $u_F^- = -\infty$ (20) holds, while in a neighborhood of $u_F^+ = \infty$ (16) holds.

Here $G(u, v) = G(-v, -u)$, and by simple calculations one can obtain that

$$\arg \min_{-\infty<u<v<\infty} G(u, v) = (-v^*, v^*),$$

where $v^* = 1.1106$ is the solution of the equation

$$\frac{1/v - 2v}{f(v)} = \frac{2 - 3F(v)}{(2F(v) - 1)(1 - F(v))}.$$

Therefore, $q^* = F(v^*) = 0.8666,\ \ p^* = 1 - q^* = 0.1334$.

The optimal confidence region has the risk of order $1/n$.

Calculations of the risk function for $(p, q) = (0.25, 0.75)$ (the first table) and for $(p^*, q^*) = (0.1334, 0.8666)$ (the second table):

n	k	m	$\lambda_2(A)$	$E_{(0,1)} t_2^2(y)$	$R(\theta, B_A)/\theta_2^2$
30	8	23	0.520142	1.869376	0.972341
40	10	31	0.338648	2.079427	0.704193
50	13	38	0.294872	1.850483	0.545655
60	15	46	0.219608	1.991273	0.437299
70	18	53	0.204000	1.841996	0.375767
80	20	61	0.166246	1.947792	0.323812
90	23	68	0.156422	1.837179	0.287375
100	25	76	0.133028	1.921895	0.255665
500	125	376	0.026954	1.840690	0.049613

n	k	m	$\lambda_2(A)$	$E_{(0,1)} t_2^2(y)$	$R(\theta, B_A)/\theta_2^2$
30	5	26	0.182234	4.335487	0.790073
40	6	35	0.114722	4.794549	0.550040
50	7	44	0.086356	5.096922	0.440149
60	9	52	0.079592	4.625095	0.368120
70	10	61	0.063934	4.855632	0.310440
80	11	70	0.053892	5.037201	0.271464
90	13	78	0.050402	4.726140	0.238206
100	14	87	0.043930	4.879758	0.214367
500	67	434	0.008606	4.952811	0.042623

Acknowledgments The authors are grateful to the referee for useful suggestions improving the paper.

Appendix

Here we establish three useful auxiliary lemmas.

Lemma 1 *Let $\{\xi_n\}$ and $\{\eta_n\}$ be two sequences of random variables and assume that there exist sequences of positive numbers $\{a_n\}$ and $\{c_n\}$ and sequences of numbers $\{b_n\}$ and $\{d_n\}$ such that the sequence of two-dimensional random vectors $(a_n(\xi_n - b_n), c_n(\eta_n - d_n))$, as $n \to \infty$, converges in distribution to a random vector with continuous density function $f(u_1, u_2)$. If $a_n \to \infty$, $c_n \to \infty$, $b_n \to b \in R$, $d_n \to d \in R$, $d \neq b$, then*

(i) under the condition $a_n = c_n$ the random vector

$$a_n\Big(\frac{b_n\eta_n - d_n\xi_n}{\eta_n - \xi_n}, \frac{d_n - b_n}{\eta_n - \xi_n} - 1\Big), \quad as\ \ n \to \infty,$$

converges in distribution to the random vector with the density $|d-b|f(-v_1 - bv_2, -v_1 - dv_2)$;

(ii) under the condition $a_n \gg c_n$ the random vector

$$\Big(a_n\frac{(b_n - \xi_n)(d_n - b_n)}{\eta_n - \xi_n}, c_n\Big(\frac{d_n - b_n}{\eta_n - \xi_n} - 1\Big)\Big), \quad as\ \ n \to \infty,$$

converges in distribution to the random vector with the density $|d - b| f(-v_1, -(d - b)v_2)$;

(iii) under the condition $a_n \ll c_n$ the random vector

$$\Big(c_n\frac{(d_n - \eta_n)(d_n - b_n)}{\eta_n - \xi_n}, a_n\Big(\frac{d_n - b_n}{\eta_n - \xi_n} - 1\Big)\Big), \quad as\ \ n \to \infty,$$

converges in distribution to the random vector with the density $|d - b| f((d - b)v_2, -v_1)$.

Proof We establish only statement (i); the other cases are treated similarly. The transformation $(u_1, u_2) \mapsto (v_1, v_2)$ of

$$a_n(\xi_n - b_n, \eta_n - d_n) \quad \text{onto} \quad T_n = a_n\Big(\frac{b_n\eta_n - d_n\xi_n}{\eta_n - \xi_n}, \frac{d_n - b_n}{\eta_n - \xi_n} - 1\Big)$$

is given by

$$\begin{cases} v_1 = \frac{a_n b_n u_2 - a_n d_n u_1}{u_2 - u_1 + a_n(d_n - b_n)} \\ \\ v_2 = \frac{a_n u_1 - a_n u_2}{u_2 - u_1 + a_n(d_n - b_n)} \end{cases} \Longrightarrow \begin{cases} u_1 = -\frac{v_1 + b_n v_2}{1 + v_2/a_n} \\ \\ u_2 = -\frac{v_1 + d_n v_2}{1 + v_2/a_n}. \end{cases}$$

This transformation has the Jakobian

$$J(v_1, v_2) = \frac{|d_n - b_n|}{(1 + v_2/a_n)^3}.$$

Therefore, the density function of the random vector T_n has the form

$$\frac{|d_n - b_n|}{(1+v_2/a_n)^3} f\Big(-\frac{v_1 + b_n v_2}{1+v_2/a_n}, -\frac{v_1 + d_n v_2}{1+v_2/a_n}\Big).$$

Since $a_n \to \infty$, $b_n \to b$, $d_n \to d$, as $n \to \infty$, statement (i) follows.

Lemma 2 *For $\beta > 0$ consider two densities*

$$f_\beta(u) = \beta(-u)^{\beta-1} e^{-(-u)^\beta}, \qquad g_{\beta,j}(u) = \beta(-u)^{j\beta-1} e^{-(-u)^\beta}/(j-1)!, \quad u < 0$$

where $j > 1$, and let (a, b) and (a', b') be such intervals that for given $\alpha \in (0, 1)$

$$\int_a^b f_\beta(u)du = \int_{a'}^{b'} g_{\beta,j}(u)du = \alpha.$$

If $\beta \le 1$, then for any $j > 1$

$$\min_{(a,b)}(b-a) \le \min_{(a',b')}(b'-a').$$

Proof Let F_β and $G_{\beta,j}$ be the distribution functions corresponding to the densities f_β and $g_{\beta,j}$, respectively, i. e.

$$F_\beta(u) = \begin{cases} \exp(-(-u)^\beta), & u < 0 \\ 1, & u \ge 0, \end{cases}$$

$$G_{\beta,j}(u) = \begin{cases} \exp(-(-u)^\beta)[1 + (-u)^\beta/1! + \ldots + (-u)^{(j-1)\beta}/(j-1)!], & u < 0 \\ 1, & u \ge 0 \end{cases}$$

and let X_β and $Y_{\beta,j}$ be the corresponding random variables.

Recall two notions of stochastic ordering (see e.g. Sect. 1A and Sect. 3B of Shaked and Shanthikumar 2007):

$X_\beta \le_{st} Y_{\beta,j}$, if $F_\beta(u) \ge G_{\beta,j}(u), \quad \forall u > 0;$
$X_\beta \le_{disp} Y_{\beta,j}$, if $F_\beta^{-1}(q) - F_\beta^{-1}(p) \le G_{\beta,j}^{-1}(q) - G_{\beta,j}^{-1}(p), \quad \forall\, 0 < p \le q < 1.$

Thus, if we show that $X_\beta \le_{disp} Y_{\beta,j}$, the lemma follows.

It is not difficult to check that $-X_1 \le_{st} -Y_{1,j}$ and that $-X_1 \le_{disp} -Y_{1,j}$ for any $j > 1$ (e.g. by Theorem 3.B.18 of Shaked and Shanthikumar 2007). Moreover, $-X_\beta$ has the same distribution as $(-X_1)^{1/\beta}$, and $-Y_{\beta,j}$ has the same distribution as $(-Y_{1,j})^{1/\beta}$. Since the function $\phi(u) = u^{1/\beta}$, $u \ge 0$, is increasing and convex, from Theorem 3.B.10 of Shaked and Shanthikumar (2007) it follows that $(-X_1)^{1/\beta} = \phi(-X_1) \le_{disp} \phi(-Y_{1,j}) = (-Y_{1,j})^{1/\beta}$. Therefore, $-X_\beta \le_{disp} -Y_{\beta,j}$, but the latter means, from Theorem 3.B.6 of Shaked and Shanthikumar (2007), that $X_\beta \le_{disp} Y_{\beta,j}$.

Numerical calculations show that Lemma 2 is true not only for $\beta \in (0, 1]$, but also for $\beta \in (1, 2)$.

Lemma 3 *If in a neighborhood of* u_F^+ *(15) with* $\beta < 2$ *holds, then for* $j \to \infty$, $j/n \to 0$, $n \to \infty$, *the relation* $nf(F^{-1}(1 - j/n))/j^{1/2} \ll (u_F^+ - F^{-1}(1 - 1/n))^{-1}$ *holds.*

Proof Let $z_n = F^{-1}(1 - 1/n)$ and $d_n = F^{-1}(1 - j/n)$. Our goal is to prove that $nf(d_n)/j^{1/2} \ll (u_F^+ - z_n)^{-1}$, $n \to \infty$. By (15) and equality $1 - F(d_n) = j/n$ it is equivalent to

$$\frac{\sqrt{j}}{u_F^+ - d_n} \ll \frac{1}{u_F^+ - z_n}, \quad n \to \infty. \tag{23}$$

To show (23) remind that for distributions satisfying (15) we have

$$1 - F(v) = L(1/(u_F^+ - v))(u_F^+ - v)^\beta \text{ as } v \uparrow u_F^+$$

(see, for example, Theorem 3.3.12 of Embrechts et al. 1997). Since

$$j = \frac{1 - F(d_n)}{1 - F(z_n)} = \frac{L(1/(u_F^+ - d_n))}{L(1/(u_F^+ - z_n))} \left(\frac{u_F^+ - d_n}{u_F^+ - z_n} \right)^\beta$$

and $j \to \infty$ as $n \to \infty$, we obtain $(u_F^+ - z_n)/(u_F^+ - d_n) \to 0$, $n \to \infty$, for $\beta > 0$. Then it is obvious that for $\beta < 2$

$$\sqrt{j}\, \frac{u_F^+ - z_n}{u_F^+ - d_n} = \sqrt{\frac{L(1/(u_F^+ - d_n))}{L(1/(u_F^+ - z_n))}} \left(\frac{u_F^+ - z_n}{u_F^+ - d_n} \right)^{1 - \beta/2} \to 0, \; n \to \infty.$$

References

Alama-Bućko M, Nagaev AV, Zaigraev A (2006) Asymptotic analysis of minimum volume confidence regions for location-scale families. Applicationes Mathematicae (Warszawa) 33:1–20

Alama-Bućko M, Zaigraev A (2006) Asymptotics of optimal confidence regions for location-scale parameters basing on two ordered statistics. In: Statistical methods of estimation and testing hypotheses. Perm University, Perm, pp 49–65 (in Russian)

Arnold BC, Balakrishnan N, Nagaraja HN (1992) A first course in order statistics. Wiley series in probability and mathematical statistics. Wiley, New York

Czarnowska A, Nagaev AV (2001) Confidence regions of minimal area for the scale-location parameter and their applications. Applicationes Mathematicae 28:125–142

David HA, Nagaraja HN (2003) Order statistics. Wiley, New York

Einmahl JHJ, Mason DM (1992) Generalized quantile processes. Ann Stat 20:1062–1078

Embrechts P, Klüppelberg C, Mikosch T (1997) Modelling extremal events for finance and insurance applications of mathematics (New York), vol 33. Springer, Berlin

Ogawa J (1998) Optimal spacing of the selected sample quantiles for the joint estimation of the location and scale parameters of a symmetric distribution. J Stat Plan Inf 70:345–360

Reiss R-D (1989) Approximate distributions of order statistics: with applications to nonparametric statistics. Springer series in statistics. Springer, New York

Shaked M, Shanthikumar JG (2007) Stochastic orders and their applications. Academic Press, San Diego